DEPARTMENT OF HEALTH & HUMAN SERVICES

National Institute of Healt
National Cancer Institut
OEM, RCB, TC
Executive Plaza South, Room 60
Phone: 496-8620 FAX 402-669

Memorandun

Date: December 22, 1997

From: Dr. Ihor J. Masnyk
Project Officer

Subject: Technical Report

To: Mr. Todd Cole TC
Contracts Specialist

Contractor: The Trustee of Columbia University

Contract number: N02-CB-77032

Reporting Period Covered: 09/30/97-12/31/97

Type of report: Quarterly Progress Report

I have reviewed the above mentioned technical progress report

✓ The report is satisfactory

The report is unsatisfactory, see comments below

COMMENTS:

Signature of Project Officer

Columbia School *of* Public Health

CSPH

Division of Epidemiology

December 19, 1997

Nancy Coleman
Contracting Officer
NCI

Dear Ms. Coleman:

Enclosed please find your copy of the First Quarterly Report for the contract entitled, "Support and Management for the Project, 'Effects of the Chernobyl Accident on Thyroid Cancer and Leukemia/Lymphoma'" (NO2-CB-77032).

Sincerely,

Sally Hodgson
Chernobyl Coordinator

Mailing Address
600 West 168th Street
New York, NY 10032

Office Location
622 West 168th Street
PH-18

Phone 212-305-9412
Fax 212-305-9413

SUPPORT AND MANAGEMENT FOR THE PROJECT: "EFFECTS OF THE CHERNOBYL ACCIDENT ON THYROID CANCER AND LEUKEMIA/LYMPHOMA" CONTRACT BETWEEN NATIONAL CANCER INSTITUTE AND THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK QUARTERLY PROGRESS REPORTS, SEPTEMBER 30, 1997 - DECEMBER 31, 1997

1. INTRODUCTION

The above contract for investigators from the Columbia University School of Public Health, investigators from other Columbia University departments and investigators from the University of Utah to provide scientific and technical support for studies on thyroid cancer and leukemia/lymphoma following the Chernobyl accident, currently being conducted in Belarus and Ukraine, came into effect on September 30, 1997. The present report covers activities by the Columbia University team of investigators for the first three months of the contract, i.e., until December 31, 1997, together with an outline of some proposed activities for the upcoming three-month period. i.e., until March 31, 1998.

The first three months of the contract have focused on the recruitment of administrative staff, and the establishment of the central facilities at Columbia through which the scientific and technical support will be managed. There have, however, been some additional activities including a visit by the principal investigator, Dr.G.R. Howe, to Kiev in October 1997 with NCI staff and consultants to discuss the leukemia project, a visit by the NCI project officer, Dr. I. Masnyk, and NCI contract specialist, Mr. Todd Cole on November 14, 1997 to meet the Columbia team and inspect the Columbia facilities, including laboratories. Finally, all relevant material has been provided to all members of the Columbia team in order to provide them an opportunity to become thoroughly acquainted with the details of the studies. Each of these items is now presented in more detail.

2. RECRUITMENT OF ADMINISTRATIVE STAFF

The establishment of the Columbia Team is complete. The Project Coordinator is Ms. Sally Hodgson, formerly Divisional Administrator of Epidemiology (see curriculum vitae in Appendix 1). She has been involved with the project since our first submission, has worked with multi-site, multi-national Former Soviet Union/United States of America studies before at Columbia University and has worked with both Professors G. Howe and J. D. Burch over the last three years. The Administrative Assistant is Dr. Svetlana Kokoreva who was recommended to us by one of our advisors, Dr. Frank Miller, Chairman of the Department of Slavic Languages of Columbia University (see letter of recommendation in Appendix 2). Dr. Kokoreva's background is Russian; however, she was born and schooled in the United States before moving to Moscow for high school and university. She has prior translation experience at the United Nations (see curriculum vitae in Appendix 1) and will be both translating and working with our desktop publishing system in addition to other duties.

3. PROVISION OF PHYSICAL FACILITIES

During the first quarter the Columbia School of Public Health designated a suite of four offices, with two open spaces intended for secretaries, in the Division of Epidemiology for the core management team and personnel occupying those offices were relocated. This suite provides contiguous space that will afford the best opportunity for easy interactions between team members. Telephone installations are complete, working file space is available and additional files and other office furniture and equipment being provided by the School are scheduled for January, 1998 delivery. Dr. Howe's office and his personal administrative assistant occupy an additional two offices adjacent to these core facilities.

4. COMPUTER AND OTHER EQUIPMENT

Five state of the art computers, a fax/copier/scanner and the necessary components of a dedicated desktop publishing system were purchased by the Columbia School of Public Health. This system is being assembled for the project by the Columbia School of Public Health systems manager who will also assist in training members of the management team in its use. Internet interface software is also being installed under the guidance of the system manager and will be tested by downloading between computers at Columbia to ensure future electronic transfer of high quality documents between the United States and Belarus and Ukraine.

Two IBM 770 Laptop computers, also provided by the Columbia School of Public Health, are scheduled for January1998 delivery. These computers, together with the Intel cameras already delivered, are capable of video conferencing through normal telephone lines and will be available for conferences as required.

5. VISIT BY DR. HOWE TO KIEV

Dr. Howe accompanied Drs. I. Masnyk, G. Beebe, S. Finch, G. Littlefield and A. McFee on a visit to Kiev from October 13 to October 18, 1997 in order to meet with those Ukranian scientists involved in the leukemia Phase I project. Drs. M. Tirmarche and P. Hubert, representing the IPSN were also present for most of the discussions.

Dr. Howe's contributions related primarily to the epidemiologic aspects of the leukemia study. He had detailed discussions with Dr. N. Gudzenko on issues of study design, implementation, setting up and maintaining appropriate databases for the study and communications. He, along with Drs. Beebe, Hubert and Tirmarche also met with Dr. Cortushin, Director of the State Chernobyl Registry. This meeting revealed some lack of information on how the database was assembled and other issues of the data quality, including the fact that files from the Ministry of Internal Affairs had not yet been added to the Chernobyl registry, despite previous assurances that this should already have taken place. The missing information on the registry and the missing data were identified as a high priority item for the Phase I study, since this will be critical to identifying and establishing the cohort for study.

Prior to the visit to Kiev, Dr. Howe provided Dr. Beebe with material (a) on possible strategies for identifying the cohort, including definitions of factors such as date of entry, date of first exposure, etc. and (b) a manuscript prepared by Dr. Howe on the topic of computerized record linkage which was written at a level where it is hoped to be of use to the Ukranian scientists involved in this process. Copies of these two documents are attached as Appendices 3 and 4.

A complete overall summary of the visit to Kiev is contained in the trip report by Dr. Beebe, dated November 8, 1997.

6. VISIT BY DR. I. MASNYK AND MR. T. COLE TO COLUMBIA UNIVERSITY, NOVEMBER 14, 1997

Dr. Masnyk, the NCI project officer, and Mr. Cole, contract specialist, visited Columbia University on November 14, 1997 as a "start-up" visit for the contract. They met with the core administrative staff (Dr. Howe, Professor Burch, Ms. Sally Hodgson, Drs. Medvedovsky, Svetlana Kokoreva, and Lydia Zablotska) for a short time, followed by more detailed discussions between Dr. Masnyk, Dr. Howe and Professor Burch on some of the scientific aspects of the studies. During this time, Mr. Cole met with Ms. Hodgson and Mr. Dias from the Columbia grants office to review the terms and conditions of the contract. This was followed by an inspection by Dr. Masnyk and Mr. Cole of the office facilities which have been provided for the management of the study.

The NCI visitors met with the other scientific investigators from Columbia (Drs. J. Fayter, D. Fink, C. Geard, E. Greenebaum, D. Heitjan, A. Matsushima, R. Reiss, B. Worgul) over lunch; Dr. McConnell, the other member of the Columbia scientific team, was traveling and unable to be present at the meeting. In the afternoon, Dr. Masnyk and Mr. Cole visited the laboratory facilities relevant both to the thyroid studies and the leukemia study, under the direction of Drs. E. Greenebaum and R. Reiss, respectively. A brief report on this visit is provided in Appendix 5.

Finally, at a debriefing session, Drs. Masnyk and Howe discussed some of the future possible involvement of the Columbia team in the studies.

7. PROPOSED WORK PLAN, JANUARY-MARCH 1998

Definitive plans have not been completed for the next three-month period, since under the terms of the contract the direction of such work will come from NCI. However, some issues can be identified where work will be required including the following:

(a) We are planning to hold an orientation workshop for the Columbia investigators in January in which we will use our advisors to present a series of seminars on the former Soviet Union, focusing, in particular, on current conditions in Belarus and Ukraine. This will alert our investigators to issues such as cultural, political and scientific sensitivities in these countries and how these may impact upon the running of the studies. Our advisors have substantial experience in this area, including frequent trips to these countries.

(b) We are anticipating that several of our investigators representing epidemiology, data processing, biostatistics and clinical disciplines relevant to the thyroid studies will participate in a visit to Belarus and Ukraine in late January/early February 1998 in conjunction with the thyroid studies in those two countries. This will provide an opportunity for the Columbia investigators to meet with NCI consultants already working in this area and with the Belarusian and Ukranian scientists involved in these studies in order to provide assistance in the solution of any problems as they relate to the thyroid studies.

(c) Similarly, we are anticipating participation from our investigators representing epidemiology, data management, biostatistics, hematology, pathology and laboratory management to take part in a visit to Ukraine to discuss aspects of the leukemia study which seems likely to take place some time in March.

(d) We are planning to do some more development work in relation to the questionnaires which will be used in the studies. For the leukemia study, this will focus on the non-dose related parts of the questionnaire already developed, and for the thyroid study, will focus on all elements. In addition, we will start planning the development of instruction manuals for interviewers and coders to accompany these questionnaires. This will of course make use of any available material including, for example, the current instruction manual being prepared at the IARC in Lyon for the leukemia questionnaire. Professor Burch will be responsible for these activities.

(e) Dr. Howe will continue to work on computer simulation studies for the leukemia study in order to assess the most appropriate and efficient study design including, in particular, strategies for selecting the sub-cohort.

(f) Dr. Howe could start developing procedures for computerized record linkage using Russian language identifiers which could be readily implemented and used in both countries as necessary; this development will be facilitated by the availability of our three Russian-speaking investigators, Drs.Kokoreva, Medvedovsky and Zablotska

(g) We are assuming that one or more candidates for training in the U.S. will be identified by NCI during this time period, and thus we will be responsible for making arrangements for such individuals including developing an appropriate schedule for training.

APPENDIX 1

CURRICULUM VITAE: MS. SALLY HODGSON AND DR. SVETLANA KOKOREVA

Sally Hodgson

19 Kevin Court, Congers, New York 10920
914-267-2629 email: sja1@columbia.edu

Education: **Michigan State University, B.A. 1967**
Michigan State University, Teaching, 1969
Columbia University General Studies, 1988

WORK EXPERIENCE:

1994 -- present Columbia School of Public Health
Division of Epidemiology
622 West 168th Street
New York, N. Y. 10032

Divisional Administrator for Epidemiology in the Columbia School of Public Health.
Reference: Geoffrey R. Howe, Professor and Head (212) 305-4601.

Responsible for administrative operations, financial accounting and academic appointments and promotions for the Division of Epidemiology. Prepare budget for central funds ($450,000) and monitor all 4-ledger budgets and all contracts and grant budgets ($7 million/year) with FSR to agency and reconciliation to the School of Public Health. Supervise compilation and reconciliation of FAS expenditure reports. Monitor centralized accounts payable systems. Maintain departmental shadow financial system and work with FFE and CAPS. Oversee grants for budget control and compliance with the Internal Audit guidelines of the University and agency requirements. Prepare various analytic projections including faculty funding and salary assignments. Support special projects as assigned. Prepare budgets for and assist with grant and contract proposal submissions for all proposals submitted through Division.

Supervise purchase orders and follow up with vendors; monitor computerized expenditure records. Prepare cash receipt deposits and track on FAS. Supervise faculty appointments and promotions providing instructions for referee letters and COAP packets. Supervise hiring of personnel, postings, advertisements and affirmative action. Maintain personnel records; interface with Union, Columbia, faculty and employees to ensure compliance with all contractual and legal obligations and to resolve conflicts and disputes.

1986 -- 1994 **Lamont-Doherty Earth Observatory**
Palisades, New York 10964

Administrative Assistant to the Oceanography Section of the Oceans and Climates Division of Lamont-Doherty Earth Observatory. Reference: Professor John Marra, Acting Director of the Observatory (914-365-8348)

Managed grants averaging $1.4 million per year from proposal development to final reconciliation: forecasted salary support for staff and scientists, monitored Principal Investigators funding levels, prepared proposals, including budgets and justifications. Organized multi-institutional and international conferences for US/USSR/FSU Joint Antarctic Weddell Sea Ice Camp. Administered visiting scientist program for both long term and short-term research and training visits of 20-30 scientists from the USSR/FSU to the US and 5-10 US scientists to USSR/FSU. Interacted with National Science Foundation personnel and other appropriate government agencies.

1984 – 1986 **Dorfman and Lynch, Attorneys at Law**
North Broadway, Nyack, New York

Para-Legal to Mr. Dennis Lynch, Esq. (914-353-3500) Introduced computerized processing for legal documents. Independently prepared separation agreements, will, mortgage riders, and other legal documents for Mr. Lynch's review. Met with clients, scheduled and prepared real estate closing materials including calculation of closing costs

1975 – 1983 **Smith Barney, Harris Upham and Co., Inc.**
1345 Avenue of the Americas
New York, New York 10019

Assistant Oil Industry and Oil Service Industry Analyst to
Kenneth S. Miller (currently Sr. VP and Managing Director, Oil and Gas Service, Lehman Bros., Inc., 3 World Financial Center, N.Y., N.Y. 10285 ((212) 298-3729)

Prepared statistical reports on small companies for the approval of Kenneth Miller and assisted in the preparation of major oil company reports and oil service industry reviews. Implemented changeover to personal computers, including advanced spreadsheets. Maintained extensive statistical records and contacted oil company senior executives to follow-up on Dow Jones reports. Prepared in depth material based on Annual Reports, 10Ks and 10Qs. Developed earnings model for the offshore oil service industry drilling fleet. Researched industry surveys through personal contacts in government agencies, publications research departments and specialized oil industry associations both in the U.S. and internationally. Organized, attended and reported in writing on meetings and receptions for Smith-Barney. Responded to inquiries from salesmen on the oil industry.

Managed office and supervised secretaries. Handled inter-office news broadcasts, correspondence and telephone contacts as requested. Handled confidential materials personally including conferences to determine action and typing of material. Monitored production of reports including duplication, printing and distribution throughout the SBHU network.

Was asked to return to work after having children on a two-three day per week office schedule with statistical work done at home. Declined Assistant Vice-President position.

1973 – 1975 **New York State Council of Environmental Advisors**
2 Wold Trade Center, New York, N.Y.

Coordinator of the Council of Environmental Advisers. Reference: Eunice Whittlesey, currently Republican National Committeewoman (518) 399-0863)

This Council to the Governor, under the direction of John L. Loeb, Jr., Chairman, was mandated by legislation in 1970 to identify environmental problems. Studies were completed and results reported direct to Gov. Nelson Rockefeller on returnable bottle legislation (including statewide public hearings), visual pollution, waste oil recovery and environmental education. "Keep New York State Clean" was an ongoing program of the Council geared to raise the level of awareness of the people of New York to the ramifications of litter; esthetically, medically and economically. This campaign coordinated the efforts of 65 county leaders and over 30,000 volunteers. Organized and maintained volunteer network. Coordinated media relations (broadcast and press), newsletters, block meetings, regional conferences and speaking engagements in New York State and Washington, D.C. Organized fund raising benefits in New York and attended same.

SVETLANA (STELLA) KOKOREVA

315 E. 68 Street, Apt. 13H
New York, NY 10021
Tel. (212) 737-3412

PROFESSIONAL EXPERIENCE:

February 1995- September 1997 ***STRATEGIC RESEARCH INSTITUTE***
(business-to-business conference organizing company) NY, NY
Administrative Assistant

- Provided administrative support to Head of Finance Division (assisted with office correspondence, travel arrangements and project coordinating for up to 10-11 national events per year).
- Performed background research for conferences and developed marketing strategies.
- Selected topics and speakers for conferences, and coordinated presentations.
- Prepared and oversaw a variety of promotional and public relations materials.
- Prepared research reports (analysis of statistics and other numerical data, etc.).

March 1993 - January 1995 ***UNITED NATIONS,* Office of Conference Services NY, NY**
Administrative Assistant

- Scheduled and coordinated meetings of UN bodies and non-governmental organizations.
- Conducted extensive correspondence, telephone and in-person consultations for up to 10-15 meetings a day.
- Drafted and edited General Assembly and Security Council reports, diplomatic notes, etc.
- Identified new terminology and verified usage in English, Russian and Spanish.
- Supervised library operations with audio/visual materials.

EDUCATION

1992 **Degree in *Political Science***
obtained from *INSTITUTE OF THE USA & CANADA STUDIES*

SKILLS

COMPUTER IBM PC & Compatibles; WINDOWS; WordPerfect 6.1, Word 6.0, Excel 5.0, PowerPoint 4.0, Access 2.0

LANGUAGES Russian (fluent)
Spanish (basic working knowledge)

REFERENCES

Available upon request.

UNITED NATIONS NATIONS UNIES

INTEROFFICE MEMORANDUM MEMORANDUM INTERIEUR

TO: A: Mrs G. Narang, Administrative Officer
Executive Office, OCS

DATE: 15 October 1993

REFERENCE: DRATS/93/113

THROUGH: S/C DE:

FROM: DE: Nigel Cassar, Chief
DRATS/TED/OCS

SUBJECT: OBJET: Evaluation of Staff Member

1. I wish to thank you for assisting staff member Ms Svetlana Tcheremouchkina who has been working in this service for the last six months. As discussed, I would strongly recommend the staff member for an administrative function. She is resourceful, helpful, punctual and has ecellent working relations with all her colleagues. She has helped to re-organize our filing system, has helped out in all capacities throughout the office and quickly learned a variety of tasks. She has helped with the Russian Terminologist, identifying new terms and verifying accuracy and usage, and helped with the job records and work distribution in the section. She has always been cheerful and willing to help regardless of pressure and time constraints. She has coped with stressful situations where a variety of problems have arisen simultaneously and has continued with her jobs to their completion. I am satisfied that the staff member could be very usefully employed in the organization.

2. I wish to thank you for your kind consideration in this matter.

cc Mr F Riesco
Ms N Tolani
Mrs. I. Bahlouli (Chron. file)

Supervisors of staff employed on a short-term appointment are requested to complete the following form in order to assist the Office of Human Resources Management in considering applications for re-employment.

NAME: Kokoreva (Last) Svetlana (First) DAM (Department) OCSS/TED. (Division) DRATS (Section)

FUNCTIONAL TITLE: Research Assistant DATE OF ENTRY ON DUTY: 12 June 1993

CATEGORY AND LEVEL: P-2 PERIOD OF SUPERVISION: 12/06/93-31/12/93

DUTIES AND ASSIGNMENT: Information functions and review of data system

Ratings on Specific Items: For each of the following items, give the staff member a rating on a 5-point scale, with 5 representing the highest possible rating and 1 the lowest possible rating.

1. TECHNICAL and PROFESSIONAL COMPETENCE [5]
2. QUALITY of WORK [5]
3. QUANTITY of WORK [5]
4. PUNCTUALITY in ATTENDANCE [5]
5. INITIATIVE [5]
6. RESPONSIBILITY [5]
7. PERSONAL RELATIONS with OTHERS [5]

Overall Rating: How proficient is this staff member in general? *(CHECK ONE)*

Outstanding [X] Above Average [ ] Good [ ] Below Average [ ] Inadequate [ ]

A. Is the staff member suited for work of a supervisory nature?

YES

B. Is the staff member more suited for some other type of work than that in which he/she is employed? If so, state the nature of work suited to his/her ability and temperament.

ADMINISTRATIVE/RESEARCH

C. Other comments.

A HARD-WORKING AND CONSCIENTIOUS STAFF MEMBER

D. Would you consider the staff member for re-employment? *(CHECK ONE)*

[X] Yes, at a higher level [ ] Yes, at the same level [ ] Yes, but only at a lower level [ ] No, not at all

Attendance: Number of days absent from office: —

Date: 21 NOV. 1994 Signature of Supervisor: N. Cassar

Date: 21 November 1994 Signature of Chief of Division: Federico Riesco, Dir, TED, C

COMMENTS OF THE OFFICE OF HUMAN RESOURCES MANAGEMENT

Date: Signature:

P.10 (9-89)-E.

A & M LOGOS INTERNATIONAL INC.

Member of American Translators Association (ATA)
Member of American Association of Collegiate Registrars and Admissions Officers (AACRAO)
Member of American Association of International Educators (NAFSA)
Member of Consular Corps of Washington, D.C.

40 Rector Street, Suite 1504, New York, N.Y. 10006 Tel: (212) 233-7061, Fax: (212) 233-7167
http://www.amlogos.com e-mail: logos@quicklink.com

EVALUATION OF DIPLOMAS

DIPLOMAS ARE EQUIVALENT TO A
U.S. DOCTOR OF PHILOSOPHY DEGREE IN POLITICAL SCIENCE

Name:	**KOKOREVA (TCHEREOMUCHKINA), Svetlana**
Ref. Number:	Ko/Ru/97963
Date:	August 11, 1997
Institution(s) Attended:	MOSCOW STATE INSTITUTE OF INTERNATIONAL RELATIONS
Country:	Russia (city of Moscow)
Year(s) of Attendance:	September 1983 - June 1988
Purpose of Evaluation:	Employment

1. The presented Diploma ПВ No.470063 was awarded in 1988 by the authority of the State Examination Board upon completion of a Program in Political Science. Admission to this program requires High School graduation and entrance examinations. Five-year study was completed at Moscow State Institute of International Relations (Moscow, Russia) in 1988. The program presents a continuous curriculum which incorporates Undergraduate and Graduate study. The Diploma, in level and intent, is equivalent to a **U.S. Master's Degree in Political Science**, with a major in International Relations, granted by a regionally accredited academic institution in the United States.

2 The Diploma КД No.069593 of Candidate of Science in Political Science (Kandidat Politicheskikh Nauk - Academic Degree in the former USSR) was awarded in 1992 by the authority of the Higher Attestation Commission of the USSR Council of Ministers. The diploma was awarded upon completion of requirements for Candidate of Political Science Degree and successful defense of a dissertation. The Diploma, in level and intent, is equivalent to a **U.S. Doctor of Philosophy Degree in Political Science**, granted by a regionally accredited academic institution in the United States.

3. This is an assessment of academic records from Russia in terms of U.S. equivalents and in accordance with Articles 1 and 5 of the "Convention on the Recognition of Studies, Diplomas and Degrees concerning Higher Education in States belonging to the European Region" (Paris, UNESCO, December 21, 1979).

4. SUMMARY: It is the judgment of A & M Logos International, Inc. that **Ms. Svetlana Kokoreva (Tchereomuchkina)** has the equivalence of a **U.S. Doctor of Philosophy Degree in Political Science**, granted by a regionally accredited academic institution in the United States.

This evaluation is based on original documents and is a true and correct evaluation to the best of our knowledge and belief.

Raisa Kaminsky

Raisa Kaminsky, Evaluator
Ph.D. in Education,
Professor of New York University,
Former Professor of Kiev University and Moscow Aviation Institute,
Author of the book *"How to Get a Higher Education in the USA"*,
Member of American Association of Collegiate Registrars and Admissions Officers,
Member of American Association of International Educators

cont. on page 2

[redacted]

COURSE & CREDIT EVALUATION REPORT

Name: **KOKOREVA (TCHEREOMUCHKINA), Svetlana**

Institution(s) Attended: MOSCOW STATE INSTITUTE OF INTERNATIONAL RELATIONS
Country: Russia (city of Moscow)
Year(s) of Attendance: September 1983 - June 1988

COURSES	U.S.Sem. Credit Hrs	U.S. Grades
History of the Communist Party of the Soviet Union	7	A
Marxist-Leninist Philosophy	10	B
Scientific Communism	3.5	Cr*
Political Economy	9	B
Basics of Communist Ethics	3.5	Cr
Theory and Practice of the Current International, Communist and Labor Movement	10	B
History of International Relations and Foreign Policy of the USSR	6	B
Theory of International Relations	3.5	Cr
History of European Countries and of America	8	A
History of the Countries of Asia and Africa	4	Cr
World History	9	B
International Law	12	A
International Private Law	9	B
State Law of the USSR and Foreign Countries	10	A
Current Aspects of Propaganda in Foreign Policy	3	Cr
Current Aspects of International Relations	3.5	Cr
Basic Theory of Information and Management	3	Cr
World Economy and Basics of Economic Analysis	6	B
International Economic Relations	8	B
Basics of Diplomacy	10	A
History of the Studied Country or Region	5	A
Economics of the Studied Country or Region	5	B
State Law of the Studied Country or Region	5	B
Foreign Policy of the Studied Country or Region	6	A
First Foreign Language	10	B
Second Foreign Language	8	B
Physical Education and Sport	3	Cr
Professional Practical Training	3	Cr
Elective Course	3	Cr
Current Issues of Economic Development of the USSR	3	Cr
State Examinations:		
Foreign Language -- Spanish	-	B
Foreign Language -- English	-	B
Basics of Scientific Communism	-	B
History of International Relations and Foreign Policy of the USSR	-	B
Thesis: " O. Torrijos and Foreign Policy of Panama"	-	A

REMARKS:* "Credit" indicates that course was completed without examination

TOTAL: 189 sem/cr. hours of Undergraduate and Graduate Study

Raisa Kaminsky

Raisa Kaminsky, Evaluator
Ph.D. in Education,
Professor of New York University,
Former Professor of Kiev University and Moscow Aviation Institute,
Author of the book *"How to Get a Higher Education in the USA"*,
Member of American Association of Collegiate Registrars and Admissions Officers,
Member of American Association of International Educators

[redacted]

Date sent: Thu, 23 Oct 1997 20:45:04 -0400 (EDT)
To: sja1@columbia.edu
From: "Frank J. Miller" <FJM6@columbia.edu>
Subject: AA position

Sally,

I just spoke with Svetlana Kokoreva, whom I know quite well, and read her the job description that you sent down to Evelyn, our DA today. Svetlana will call you tomorrow morning regarding this. I think she would be a fine candidate for this position; she is Russian and grew up in the Soviet Union, but she is an American citizen - her father was working in the States when she was born and at that time they let the children of certain Soviets in the States have American birth certificates. She also has the necessary office skills, and she has done some work at the United Nations along the way. She (or her resume) can tell you more about her than I can, but if necessary, I'll be happy to write a recommendation for her.

And congratulations on the grant. I take it that everything is now official since I saw a nice article about it in the "Record" the other day.

Frank Miller

****

Frank J. Miller, Chair
Department of Slavic Languages
Hamilton 708 Mail Code 2851
Columbia University New York, NY 10027

Telephone: 212-854-7499
Fax: 212-854-5009
e-mail: fjm6@columbia.edu

****

APPENDIX 2

LETTER OF RECOMMENDATION FOR DR. SVETLANA KOKOREVA

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
DEPARTMENT OF SLAVIC LANGUAGES

December 15, 1997

Ms. Sally Hodgson
Project Director
Division of Epidemiology
School of Public Health
600 West 168th St.
New York, NY 10032

Dear Ms. Hodgson:

I am pleased to recommend Svetlana Kokoreva for an administrative assistant position in your division. I first met Ms. Kokoreva approximately six years ago when she came to New York from Moscow to visit her parents. At that time her father was part of the Soviet delegation to the United Nations, where he headed the interpretation division.

Ms. Kokoreva has a rather unique background which makes her an ideal candidate for this position. She was born in the US of Russian-speaking parents and learned English as a child. I believe she also began elementary school in the States. She received her higher education in the former Soviet Union where she also continued her study of English. Therefore, she is completely bilingual in English and Russian. Since she has American citizenship, she chose to live in the States when the Soviet Union collapsed at the end of 1991.

For the past five years Ms. Kokoreva has held various clerical jobs and has taken various courses to perfect her computer skills. With the proper fonts, our computers can also print Russian, Ukrainian, Belorussian, and Ms. Kokoreva can definitely do any work in Russian on a computer.

Ms. Kokoreva also has the cultural background necessary to work with people from the former Soviet Union. She has lived most of her life Russia, she is a person who will interact well with the people you will be dealing with and convey to them our commitment to this project.

This is a position that will utilize all of Ms. Kokoreva skills, and I recommend her most highly.

Sincerely,

Frank J. Miller

Frank J. Miller,
Departmental Chair

708 Hamilton Hall Mail Code 2839 1130 Amsterdam Avenue New York, NY 10027 212-854-3941 Fax 212-854-5009

APPENDIX 3

POSSIBLE STRATEGIES FOR IDENTIFYING THE COHORT FOR THE STUDY OF LEUKEMIA AND LYMPHOMA IN UKRANIAN CLEAN UP WORKERS

Memorandum

To: Dr. Gil Beebe

From: Dr. G.R. Howe

Date: October 9, 1997

Re: The Study of Leukemia and Lymphoma in Ukrainian Clean up Workers

---

1. Introduction

As you know, I drafted an outline of the design for this study which now constitutes Appendix 6 in the current version of the protocol. This memo addresses some further issues in this context, in particular (a) the specific definition of the underlying cohort(s) (b) the sampling strategy for the sub-cohort and (c) the most appropriate way to incorporate the retrospective cases i.e. those diagnosed between 1986 and 1997 into the study.

On reflection it seems to me that we could usefully carry out two distinct studies. The first would be a full cohort study covering the whole of the Ukraine and utilizing the cases identified to date through the emergency form mechanism. The way in which this could be carried out, and its advantages and limitations are discussed below in section 2. The other study would of course be the definitive case cohort study as outlined in Appendix 6 of the protocol. This is discussed further in section 3 below.

2. The Full Cohort Approach

This study would make use of the case data collected to date through the emergency form mechanism. The study design is as follows:

(a) It would comprise a full cohort analysis.

(b) The cohort would consist of all clean-up workers on the Chernobyl registry who were resident in Ukraine immediately after leaving the 30 km zone for the first time..

(c)Time at entry into the study will be the date of leaving the 30 km zone for the first time.

(d) Time at first exposure would be the first date non-zero dose was recorded in the Chernobyl registry.

(e) Time at exit from the study would be date of last medical exam plus one year: the latter period allows for cases diagnosed within that year, since presumably after diagnosis they are unlikely to return for the routine medical examination. An alternative definition of time at exit would be date of death or out migration from Ukraine if this could be determined by record linkage to existing data bases.

(f) Cases would consist of those identified through the emergency form mechanism, and would include both those identified to date and any identified subsequently if this mechanism remains in place. Completeness of identification of cases could be checked eventually by record linkage to the Ukrainian cancer registry when this becomes feasible.

(g) Doses will be those contained in the Chernobyl registry. However, it might be possible to calibrate these recorded doses using the corresponding doses for the sub-cohort (see below) in conjunction with biological dosimetry and/or dose reconstruction carried out for members of the sub-cohort. Unfortunately, however, this may not be feasible since the corresponding error structure is likely to be extremely complex.

The advantages in carrying out such an analysis are as follows:

(a) it would make maximum use of the existing data i.e. cases identified through the emergency forms.

(b) it would require no extra data collection.

(c) if the linkages mentioned above prove to be feasible, they will have to be developed for the definitive study in any case. Using these same linkages for more records does not require extra resources.

(d) the proposed analysis could be done relatively soon.

(e) the analysis could provide a model for the statistical approach to be used for the definitive case cohort study.

(f) the analysis could provide useful experience for our Ukrainian colleagues; analysis could be conducted as part of the training program.

(g) if it is intended to continue collecting data for the whole Ukraine from the emergency medical forms, in the future, this proposed type of analysis would seem to be the sensible way of using those data.

There are of course severe limitations to the full cohort study as follows:

(a) there would be no definitive diagnostic review of the cases to confirm the diagnosis.

(b) the dosimetry is likely to prove inadequate.

(c) there is no guarantee that identification of cases is complete, until and if the cancer registry data become fully available.

(d) data on other covariates may not be available for study subjects.

Despite these limitations, it might be worthwhile considering such an exercise, particularly since it only requires limited resources, would prove useful for training, and probably makes the best use of existing data.

3. The Definitive Case Cohort Study

This is the study outlined in Appendix Six of the existing protocol. The present material attempts to provide more specific definitions of the study cohort and sub cohort and in addition to address the issue of how to incorporate data on retrospective cases. The fundamental approach to the latter question is to define the sub cohort in such a way that it provides a valid comparison group for both retrospective and prospective cases. The suggested design is as follows:

(a) The underlying cohort is defined as all clean-up workers on the registry first entering the 30 km zone between 1986 and 1990 who subsequently returned to or took up residence in the six oblasts. There should be no limitation on the time at which the latter residence was started.

(b) Time at entry into the study is defined as the first date returning to or taking up residence in the six oblasts following the date of first entry to the 30 km zone.

(c) Time at first exposure will be defined by the dosimetry study.

(d) Time at exit from the study could be the last date of medical examination in which the subject was still resident in the six oblasts plus 1 year (for reasons discussed above). Alternatively, if feasible, record linkage could be used to identify fact and date of death, or fact and date of migration from the six oblasts if suitable data bases exist. In addition, it may be feasible to consider active follow-up of the sub-cohort to determine date of out migration from the six oblasts, or date of death. The choices amongst these methods will depend on the information collected during phase 1.

(e) Cases will be identified prospectively by surveillance at the oblast cancer hospitals, supplemented by linkage between the hospital records and the Chernobyl registry. Retrospective cases will be identified by linkage between the corresponding retrospective hospital records and the Chernobyl registry.

(f) All cases will be subjected to a thorough diagnostic review with all available material.

(g) The sub-cohort will be a random sample of the full cohort as defined above. However, it may be stratified by factors such as oblast of residence, age and year of first entry into the 30 km zone. The size of the sub-cohort and stratification variables are currently being investigated by computer simulation which is an extension of that described in Appendix Six of the existing protocol.

(h) Dosimetry and other covariate data will be obtained for cases and all members of the sub-cohort.

The above approach should provide a valid comparison group for both retrospective and prospective cases. However, it is necessary to take into account the fact that the dosimetry data for retrospective cases is likely to be of a poorer quality than that for the sub-cohort, in that biological dosimetry would not be possible for such cases, and any dose reconstruction information is likely to be by proxy. The impact of this i.e. greater measurement error for cases than for the sub-cohort, depends upon the error structures involved. For example, as an alternative one could consider defining a sub-cohort with dosimetry collected in the same way as for retrospective cases i.e. with greater measurement error. However, such a sub-cohort would of course be inappropriate for comparison with prospective cases. Further, if measurement error is uncorrelated with exposure, less bias is introduced by reducing measurement error for the sub-cohort even when the cases have greater measurement error. Therefore, it is recommended that the subcohort (a) be selected from the full cohort defined as above from the time of first residence in the six oblasts following entry to the 30 km zone and, (b) have as full and accurate dosimetry performed on them as possible.

Several other points are worth noting. It is not necessary to choose the full sub-cohort at the beginning of phase 1; since the objective is only to obtain a representative sample for the pilot work this could be done independently of selecting the sub-cohort, or alternatively the pilot sample could eventually form part of the sub-cohort. It may be sensible to wait until any pilot results are available before finally defining the criteria for the sub-cohort.

Secondly, there may be other restrictions on cohort membership such as incompleteness of data in the registry, or failure to pass consistency checks on the registry data.

Finally, detailed tabulation of the registry data are an essential first step before completing the specifications for cohort definition and sub-cohort sampling.

cc: Ihor Masnyk, Ph.D.

Bruce Wacholz, Ph.D.

APPENDIX 4

THE USE OF COMPUTERIZED RECORD LINKAGE IN COHORT STUDIES

GEOFFREY R. HOWE

October 14, 1997

The Use of Computerized Record Linkage in Cohort Studies

Geoffrey R. Howe[1]

[1]Professor and Division Head
Division of Epidemiology
Columbia University School of Public Health
622 West 168th Street, Room PH18-119
New York, NY 10032

INTRODUCTION

The term "Record Linkage" refers to the process of comparing two or more records which contain identifying information, to determine whether those records refer to the same entity. In the present context, the entity in question will normally be an individual enrolled in a cohort study. It would be more accurate to say that the objective of a linkage is to estimate the probability that the records refer to the same person, since in practice one can never be 100% certain that indeed this is the case. The two fundamental problems which so limit the interpretation of record linkage are duplication of identifiers, and errors in creating or transmitting records. These problems are further considered below.

Record linkage as such has a long history in epidemiology and many other disciplines. However, whereas the process used to be limited to a manual one, the advent of high speed computers has radically changed our ability to carry out record linkages on a massive scale involving many thousands or indeed millions of records. Although the fundamental principles remain the same, the necessity to take into account the characteristics of the computer systems involved has led to the introduction of the term "Computerized Record Linkage", the topic of the present review.

To date by far the greatest use of computerized record linkage in cohort studies has been in the context of passive follow-up of cohorts to determine factors such as vital status, residential status or the health outcomes being studied in the cohort. In passive, as opposed to active follow-up individuals in the cohort are not directly contacted, but rather are monitored using a registry of those events of interest such as deaths or cancer incidence. Such passive follow up has often made use of some identification number, for example, the Social Security number, to match cohort records to outcome records, this number being regarded as unique to a particular individual.

This illustrates an important point regarding record linkage. Although one may think of the Social Security number as a "unique identifier", in reality this will not always be so. It is conceivable that two individuals could have been issued with the same duplicate Social Security number in error, and secondly, and more importantly, when such numbers are being recorded errors will almost certainly occasionally be made. Thus, even with what is thought of as a unique identifier, errors in linkage can occur even though such errors are likely to be small.

Thus, even with unique identifiers, linkage is inevitably probabalistic in nature, and this needs to be borne in mind by researchers conducting record linkages based on such a unique identifier. In practice of course even when a unique identifier is available confirmation will be sought by comparing other identifiers on the two records e.g. name, date of birth, etc. The obvious question is then what weight to give to the possible combination of agreements and disagreements between the two records amongst the various identifiers, including the so called unique identifier. Clearly it is desirable that this process be as quantitative as possible, and the application of the rules of classical probability theory have led to the development of the theory of probabilistic record linkage (1-3).

In many situations, a unique identifier will not be available. As an example consider the Canadian fluoroscopy cohort study (4-6). In this study a total of 110,088 records were collected from forty-two institutions in Canada which had treated tuberculosis in the 1930's and 40's. Each record contained information from the original institutional records on a tuberculosis (TB) patient who had been treated in that institution. The extracted information included identifiers such as name and date of birth, and a history of treatment for tuberculosis. Some 40% of TB patients during that era were treated with a series of pneumothorax or pneumopiritoneum, during which they received substantial exposure to low linear energy transfer ionizing radiation, particularly to organs such as the breast and lung. The objective of the study was to conduct a long term follow-up to ascertain mortality and cancer incidence in the cohort, in order to relate the estimated organ specific radiation doses to subsequent risk of cancer (5-6).

Two problems in the study were addresssed by computerized probabilistic record linkage. First, because of the possibility of admissions of a single TB patient to multiple institutions, it was necessary to identify which of the 110,088 records referred to the same individual in order to establish a cohort of individuals. This process is sometimes referred to as "internal record linkage" ie comparing all the records in a single file to determine which groups of records might refer to the same individual. The second process involving record linkage was in the context of passive follow-up. As described subsequently, Canada has a national mortality data base which contains records of all those who have died in Canada since 1940. In order, therefore, to conduct the passive follow-up, the records of the cohort of TB patients were linked to the mortality records (currently between 1940 and 1987).

No identification number was available to link the records for either of these linkages. Instead, both linkages made use of identifiers such as given names, surname and date of birth. Clearly, such identifiers are much more prone to duplication than than an identification number. Hence, the use of probabalistic techniques was essential for the linkages to be carried out.

In this review a brief description is first provided of the principles involved in probabilistic record linkage, followed by consideration of some of the issues involved in the computerized application of those principles. In order to illustrate such applications in a real life context, a description is then given of the linkage system and data bases which have been developed in Canada over the past 30 years, and in particular these are illustrated by reference to a number of cohort studies which have been conducted. Consideration is then given to the US situation, in particular the US National Death Index which currently is extensively used for passive follow-up in the United States. Finally, some of the problems of record linkage and some of its potential future uses in cohort studies are discussed.

PRINCIPLES OF PROBABILISTIC RECORD LINKAGE

The theory of probabilistic record linkage has been developed by a number of authors (see for example, 1-3,7). In this section a brief overview of basic principles is given, which should be adequate to provide the necessary grounding in these principles. However, for actual implementation the reader is referred to a monograph by Newcombe (3) which provides a more detailed and practically oriented description of the techniques involved.

Consider two files of records, file A and file B containing respectively n_A and n_B records. Each record contains information on a standard set of identifiers such as surname, given names, day, month and year of birth. Conceptually one wishes to compare all records on file A with all records on file B and estimate the probability that any pair of records refer to the same individual, conditional on the identifying information contained in each record. We refer to comparisons which result in a tentative judgement that the records refer to the same individual as links or potential links; when two records truly refer to the same individual they are called matched.

When any two records are compared, the value of each identifier on each record is compared to the value of the corresponding identifier on the other record, yielding an outcome. For example, an outcome which might follow from comparing surnames is "the two records agree on the value Smith". An alternative outcome might simply be that "the records disagree on the value of the surname" i.e. without specifying what those values are.

In order to quantify the evidence provided by the various outcomes, the two key parameters are the probability of the outcome in the matched set of comparisons (P|M)and the probability of that outcome in the unmatched set (P|U). Generally, these quantities will not be known given that the purpose of the linkage is to establish the matched and unmatched sets. However, various techniques can be used to estimate these probabilities. Such estimates need to take into account the problems of both duplication and recording errors.

As a simple example consider the outcome "agreement on the surname Smith". The distribution of names on the matched set should be similar to the distribution in the cohort, assuming that the value of one's surname does not affect one's mortality risk. Hence, the probability of the outcome "agreement on Smith" in the matched set can simply be estimated from the frequency of occurrence of the name Smith in the cohort file (assuming no recording errors).

The unmatched set will of course consist of all the n_A n_B comparisons minus those comparisons which are matched; however, the latter number will be negligible compared to the former number. The maximum number of matches will be n_A if file A is the cohort assuming the mortality file is larger than the cohort file which it almost certainly will be. Thus, the ratio of non matches to matches has a maximum value of n_B which is likely to be in the many thousands or even millions. Hence, one can estimate the probability of the outcome in the unmatched set simply by multiplying the relative frequency of Smith on file A by the corresponding relative frequency on file B, since agreement by chance is simply a random phenomenon.

Thus, the ratio (P|M)/(P|U) may be estimated by:

$$\frac{(P|M)}{(P|U)} = \frac{f_A}{f_A f_B} = \frac{1}{f_B} \qquad (1)$$

where f_A and f_B are the relative frequencies of Smith on the two files respectively.

The importance of the quantity in equation 1 is that it gives a numeric measure of how much evidence the occurrence of a particular outcome provides that the two records in question do indeed refer to the same person. It is conventional, as in information theory, to take the log of this quantity to the base two: this quantity i.e. $\log_2 (P|M)(P|U)$ is referred to as the weight corresponding to that particular outcome; weights from a series of outcomes observed from the comparison of a pair of records may be added to give a total weight as they are on a logarithmic scale (see below).

This derivation of the weight corresponding to the outcome "agreement on Smith" does not take into account errors which can occur in recording data. Dealing with errors in this context depends on the error structure in the particular files being compared. Consider a situation in which 10 percent of the records in the above example have the surname mis-recorded, and in which this error rate is the same for both file A and file B, and is also independent of the value of surname (which in practice it almost certainly will not be). In our example, the frequency of Smith on the matched set will be reduced by 19 percent i.e. 10 percent + 10 percent - 10 percent. 10 percent (according to standard probability theory). On the other hand, amongst the unmatched set, if errors in recording are such that the overall distribution of names on the two files remains the same e.g. a Smith is recorded as Jones, and a Brown is recorded as Smith, the distribution of names in the unmatched set remains the same. Hence, the weight for the outcome "agreement on Smith" has the value: $\log_2 0.81 - \log_2 f_B$ where we have now taken both duplication and recording errors into account.

In practice, the true error structure is unlikely to be known, and more importantly the value of the error rate is also unlikely to be known. One possible approach to estimating the error rate (2) is described subsequently.

Similar approaches to that described above may be used to estimate weights for more complex outcomes e.g. partial agreement on surname, years of birth disagreeing by one year etc. The principles involved remain the same: Newcombe (3) has provided more detailed descriptions of the derivations of weights under a number of circumstances, including some purely empirical approaches involving random sampling to estimate frequencies of outcomes in the unmatched set.

As mentioned above, weights may be combined additively for a series of outcomes provided that these outcomes are statistically independent. If they are, the difference in weights between two comparisons [e.g. record 1 on file A with record 1 on file B, and record 2 on file A with record 2 on file B] represents the log of the odds ratio in favor of a match between those two comparisons. This may be seen from Bayes theorem since:

$$\frac{(P_M|O_1O_2\ldots)}{(P_U|O_1O_2\ldots)} = \frac{(P_{O_1}|M)}{(P_{O_1}|U)} \bullet \frac{(P_{O_2}|M)}{(P_{O_2}|U)} \bullet\bullet\bullet \frac{P_M}{P_U} \qquad (2)$$

where the quantity on the left hand side of the equation is the odds in favor of a match given the observed outcomes O_1, O_2, etc., the terms on the right are defined in equation 1 and P_M and P_U are the overall probability of a match and the overall probability of a non-match respectively.

It is clear that this could be converted to an absolute odds if one knows the value of the ratio of the number of matches to nonmatches. However, this is rarely known in advance, though rough estimates can sometimes be made e.g. using life table techniques to predict mortality in a cohort.

However, in practice most linkages are conducted in relative terms i.e. potential links can be ordered by their total weight, and this should represent the order of increasing believability of such links.

This leaves the obvious problem of where to set a cut-off or threshold value i.e. how to establish a weight above which one "believes" the links, and below which one "disbelieves" the links. This issue has been addressed theoretically by Felligi & Sunter (7), but in practice the great majority of linkages depend on empirical evaluation to set the cut-off value. Generally, if one plots the frequency distribution of weight one obtains a curve with a valley which separates the matches from the non-matches. This valley can be fairly broad, particularly if the quantity or quality of identifying information on the two files is poor, and under these circumstances, one can only "guesstimate". An alternative approach is to set two threshold values, one lower and one upper and to inspect each link between the two in an attempt to validate that link or otherwise. The latter can generally only sensibly be done when additional identifying information is available in records which were not available for the computer match.

COMPUTER PROCESSING FOR RECORD LINKAGE.

The above principles can readily be applied manually if the number of records in each file is relatively small. However, in many cases, the numbers are simply too large for such manual linkage. For example, in a study of a 10% sample of the Canadian labor force, Lindsay et al (8) linked approximately seven hundred thousand records of an occupational cohort to three million mortality records. Linkages of this size are however beyond the capacity of most computers. Hence, the concept of "blocking" or "pocketing" records has been introduced. Records are divided into such blocks or pockets on the basis of some identifying characteristic which is such that it is concluded that it is extremely unlikely that any two records which do not share this same characteristic will be matched. Comparisons are then restricted to records within a pocket.

A common practice is to use an alpha numeric code of the surname, the New York State Identification and Intelligence System (NYSIIS) code (9) to block records in this way, since it has been shown that matches very rarely occur between records with differing values of this code. Alternative pockets may be defined to check this assumption, or smaller pockets may be created by combining the NYSIIS code with, for example, gender. In the linkage only records in the same pocket are then compared. Whereas this process saves substantial computer resources, it does of course increase the possibility of false negatives i.e. missed matches.

A number of systems have been developed to make use of the type of probabilistic record linkage described above, and have been implemented on a variety of computer systems ranging from personal computers to mainframe machines. Some of these packages are commercially available. However, the overall process is not complicated, and if one does not have access to an existing system, it is relatively straightforward to program any particular application in a language such as Fortran or C.

To illustrate a typical processing system, figure 1 shows a flow chart representing the process used by the generalized record linkage system developed conjointly by the present author and staff at Statistics Canada, an agency of the Canadian federal government (2). In brief, the files are first edited to ensure appropriate data quality, and are then grouped into pockets typically defined by NYSIIS code. Each record in a particular pocket on one file is then compared to all records in the same pocket of the other file using initial weights based on frequency distributions (or alternatively including initial estimates of error rates from some external source). The preliminary matches are then tabulated to produce a (revised) error rate and the linkage is repeated with new error rates being computed until self consistency is achieved. The final links are then inspected and lower and upper thresholds are set empirically if there is additional identifying information which is not available to the computer. The links in this "gray area" are then resolved (manual resolution). Alternatively, a single threshold is set empirically if no further identifiers are available.

Finally, all links involving the same record are grouped and checked for possible conflicts, for example, one cohort record linking to two death records. Multiple linkages of the same record may of course be permissable, e.g. the internal linkage of the Canadian TB patient file described earlier.

APPLICATION OF COMPUTERIZED RECORD LINKAGE TO COHORT STUDIES IN CANADA

Starting in the early 1970s', the National Cancer Institute of Canada in conjunction with Statistics Canada initiated work on several cohort studies which depended for follow-up on computerized record linkage to national registries. The Canadian work stemmed from the seminal influence of Dr. Howard Newcombe, a population genetisist who developed much of the early theory, and inspired a number of the early epidemiologic studies which took advantage of those methods (1). Because of this work, the Canadian experience in probabilistic computerized record linkage applied particularly to health studies is probably greater than in any other country in the world. The early work involved both establishing the national mortality data in a data base form suitable for record linkage, and developing a general system for conducting such linkages.

Subsequent to the completion of the mortality data base which has records from 1940 on, Statistics Canada developed other data bases for use in health studies including the national cancer incidence data base containing records from 1969 on.

The generalized record linkage system (GRLS) developed to take advantage of these data bases is now well established and has been summarized in the previous section. It has been used to conduct many health related studies, and in particular has proved useful for cohort studies in which the end point of interest is cause of death.

Brief descriptions of several cohort studies conducted using the GRLS follow for exemplary purposes. The focus of these descriptions is on the epidemiology of the studies, since the main objective is to demonstrate the utility of the combination of computerized record linkage with passive follow-up to provide useful results in cohort studies.

(a) The Fluoroscopy Study (4-6):

This study has been described in the introduction. Its primary purpose is to assess the nature of the relationship between exposure to low LET ionizing radiation, and risks of breast, lung and possibly other cancers. Following the collection of the 110,088 records from the original institutions, an internal linkage was carried out using the GRLS to identify multiple records referring to the same individual. This process yielded a cohort of 93,192 individuals. A second record linkage was then carried out linking the composite records for cohort members to the Canadian mortality data base between 1940 and 1987. Fact of death was available between 1940 and 1949, but not cause of death which was only added starting in 1950. Thus, the final cohort for statistical analysis consisted of 64,172 subjects, 32,255 men and 31,917 women who were known to be alive at the start of 1950.

Analyses from this study with respect to breast cancer mortality (6) and lung cancer mortality (5) have been published. In addition, a third linkage has been conducted to the national cancer incidence data base, following which results with respect to breast cancer incidence have been reported (10). In brief, a strong linear dose response between breast tissue dose and breast cancer risk was seen in the cohort, with the risk per unit of dose decreasing with increasing age at exposure. A comparative analysis between the fluorospy study and the atomic bomb survivors study (6) demonstrated that the results from the two series were similar, though it has been suggested that the fluoroscopy data also provide some evidence of a dose rate effect i.e. reduction in risk per unit dose with increasing fractionation. (10)

In contrast, a comparative analysis between the two series with respect to lung cancer mortality showed dramatic differences, with no evidence of any radiation related lung cancer risk in the fluoroscopy cohort based on a total of more than 1000 lung cancer deaths, which contrasted with a highly significant risk seen in the corresponding data for the atomic bomb survivors study (5). The author has argued that this provides strong evidence of a substantial dose rate effect for lung cancer induced by low LET radiation.

This example provides a clear illustration of the benefit that can be obtained by using record linkage in cohort studies. Study subjects could have been lost to contact anytime from 1940 on, and indeed a substantial number have no follow-up information in the original records beyond their date of discharge from the institution which primarily was in the 1940's. Hence, when this study was initiated some thirty years later, active follow-up would have been essentially impossible. Thus, this study was only feasible because of the existence of the mortality data base, and the capacity for computerized probabalistic record linkage which was developed in part for the purpose of conducting this study.

(b) The Labor Force Study (8,11-12): the purpose of this study, briefly mentioned above, was to identify patterns of mortality in a 10 percent sample of the Canadian labor force and relate these to previous occupational history. The study primarily aimed to develop hypotheses which could be tested later in more definitive studies, but the data also serve to test hypotheses generated from other sources.

In brief, the sample was collected between 1965 and 1971 by the Dominion Bureau of Statistics, the predecessor of Statistics Canada. For each individual in the sample information was collected on the occupation and industry in which that individual was working in each year of the survey (the data for 1970 were subsequently lost). The epidemiologic study was initiated by converting the original punched card records to magnetic tape. The primary identifier available was the Canadian social insurance number (SIN), a so called unique identifier, but subject to the limitations previously described. In addition, limited information was available on surname, gender and year of birth. Because the SIN is generally not available on the mortality data, and because the remaining information was inadequate for accurate linkage, the first step was to add much more detailed identifying information from the master index file of SINs. Subsequently, using the much more detailed identifying information, a computerized record linkage was carried out between the seven hundred thousand records from the labor force sample and approximately 2 million death records for the years 1965 to 1973 (11). Subsequently, two updatings of the mortality linkage have been conducted.

Analyses were carried out relating the various occupational codes for individuals on the file to risk of 67 different causes of death. Because of the monitoring nature of the exercise, and the obvious problem of multiple comparisons, results have been reported for associations satisfying criteria based both on significance and size of effect. A series of publications has included the various updates, and results for both men and women (8, 11-12).

This study provides an example where active follow-up could possibly have been utilized since the cohort was established much more recently than the fluoroscopy cohort; however the cost of such follow-up would be completely prohibitive given the large size of the cohort, and hence the use of computerized record linkage with the mortality data base provides a very cost effective and indeed the only sensible approach to carrying out such a study.

(c)The Eldorado Study (13-16): this is a cohort study of workers at the company then called Eldorado Resources Limited which operated two uranium mines and a processing plant in Canada. The objective was to study the relationship between exposure to radon and its decay products and the subsequent risk of cancer, particularly lung cancer. The cohort of approximately 20,000 individuals was established from the personnel and payroll records of the company. The identifying information available was somewhat limited, as was typical of many mines at the time the two Eldorado mines were operating. However, it proved adequate for the purposes of record linkage, which was used to identify mortality in the cohort from 1940 on as in the fluoroscopy study. Again analysis had to be restricted to the years from 1950 on when underlying causes of death became available on the data base.

Separate analyses have been reported for the two mines. The first mine which was operated at Port Radium in the Northwest Territories opened in 1930 and closed in 1960. The second, Beaverlodge in Saskatchewan, opened in 1949 and closed in 1982. Both cohorts of miners demonstrated an exposure related substantial increased risk of lung cancer mortality between 1950 and 1980, the most current year for which follow up data were available at the time results were reported. However, the risk per unit of exposure differed by a factor of about 12 between the two mines, with Beaverlodge having the higher risk. It was postulated that this might be due to an inverse exposure rate effect, since exposure rates were far higher at Port Radium than they were at Beaverlodge.

Subsequently, the data from these two cohorts were included in a combined analysis with those from nine other cohorts of underground miners exposed to radon decay products, and this combined analysis provided strong evidence in favor of the inverse exposure rate hypothesis (17). The data from these cohorts have also been analyzed for cancer mortality other than lung cancer and these results too have been reported (18).

Again, the Eldorado study illustrates a situation in which study subjects may have been lost to contact many years ago, particularly those employed at the Port Radium, mine; hence, computerized record linkage with passive follow-up provides the only feasible means of conducting such a study. This study also demonstrated that the record linkage system employed is relatively robust to identifying data quantity and quality, which were poorer in the study than in most of the other cohort studies.

(d) The Canadian National Railway Study (19):The objective of this cohort study was to ascertain mortality patterns amongst retired railway workers, and in particular to determine whether relationships existed between various causes of death and occupational exposure to diesel fumes and coal dust. The cohort consisted of all male Canadian National Railway pensioners who had retired before 1965 and who were known to be alive at the start of that year as well as those who retired between 1965 and 1977, in total 43,826 subjects. Computerized record linkage was used to link records from the cohort to the Canadian national mortality data base to the end of 1977. This resulted in a total of 17,838 deaths between 1965 and 1977, of which 933 were ascribed to lung cancer as the underlying cause.

It is of interest to note that the original proposal to ascertain deaths in the cohort was for purely administrative purposes in order to carry out a death clearance of the company's pension file. However, this proposal was refused on the grounds of confidentiality, but permission was given to carry out the linkage in the context of an epidemiologic study, but with no administrative uses of the linkage permitted.

Exposure to coal dust and exposure to diesel fumes were assessed on the basis of last recorded occupation in the pension file (the only occupation available) by industrial hygienists from the railway company, who graded the occupations into "non, possibly or probably exposed". On the basis of this exposure measure there was a highly statistically significant dose response relationship between exposure to both coal dust and diesel fumes and risk of lung cancer mortality with relative risks of 1.0, 1.2 and 1.4, $p<0.001$. However, it was not possible to distinguish effects of coal dust from diesel fumes since these exposures were highly correlated. Although it was not possible to account for the potential confounding effect of cigarette smoking, these results provided some evidence in support of the postulated causal relationship between exposure to diesel fumes and lung cancer risk.

This study again illustrates the utility of computerized record linkage in carrying out large scale cohort studies at a relatively low cost. It is of interest to note that, although the cohort was based on pension records, clearly active follow-up would have lead to missing information, since such missing information was the rationale for the original proposed administrative use of the linkage.

U.S. NATIONAL DEATH INDEX

The most widespread application of computerized record linkage in the context of cohort studies in the United States is provided by the National Death Index [NDI]. The NDI is a registry of all deaths occuring in the United States, and is maintained by the National Center for Health Statistics (20). Computerized mortality records are provided to the NDI by the Vital Statistics offices in all 50 states, the District of Columbia, Puerto Rico and the Virgin Islands, generally within 12 months of the completion of each calendar year. The index contains deaths from 1979 on, with approximately 2 million deaths being added each year.

The NDI may be used by epidemiologists and others to link data from cohorts to the Index in order intially to identify date, fact and State of death for individuals in the cohort who have died. However, unlike the corresponding Canadian system described above, the linkage process to the NDI is not probabilistic in nature. Rather, all records in the two files, i.e. the cohort file and the NDI are compared on a series of definite criteria which have to be completely safisfied for any pair of records to be identified as a potential link. These criteria consist of combinations of various identifiers such as the social security number, surname, date of birth and gender. A pair of records is considered a potential match when there is exact agreement on all the identifiers contained in any particular criterion. As with the Canadian system, the NYSIIS code of surname can be employed in the NDI linkage system to allow for common mispellings of surnames.

The researcher receives back from the NDI records of all possible matches i.e. all those satisfying any of the various matching criteria. Because of the possibility and indeed probability of multiple links including the same record(s) possible links have to be resolved in some way (see below) in order to identify those links which will be treated as "definite", and those which will be rejected. The researcher may then write to the appropriate State local office in order to obtain a copy of the death certificate corresponding to the particular death link, in order to obtain the underlying and other causes of death. Recently, the NDI has announced the formation of the National Death Index Plus whereby researchers will automatically be provided with the causes of death coded to the Ninth Revision of the International Classification of Diseases for any potential link identified in the NDI, thus saving the researcher the necessity of obtaining copies of the original death certificate from the State in question.

Clearly, the major potential problem with using the NDI is that inevitably the linkage is liable to initially produce many false positive links since it is deliberately designed to err on the conservative side in the sense of identifying any potential link by virtue of its use of multiple linkage criteria. This problem is clearly recognized by the custodians of the NDI who warn researchers of this issue, and advise them of the necessity to further prune the potential links to reduce these to a more reasonable number.

Such pruning is best achieved by the use of probabalistic techniques identical to those described earlier in this review. Estimates can readily be made of the weights corresponding to the various possible outcomes, as can estimates of error rates in recording. When this has been done, a total weight may be computed for each potential link.

The next step is to resolve any duplication i.e. multiple potential links involving the same records. The simplest approach is to allow the highest weight relating to any specific record to prevail (though theoretically this does leave a small possibility of some logical conflicts). Since in general no further identifying information will be available for records from the NDI, resolution should be based on the probability weights rather than manual inspection. The final decison is again to establish a cut-off value dividing links into accepted or rejected.

The combination of the resources of the NDI together with subsequent probabalistic record linkage provides an excellent resource for conducting cohort studies in the U. S. in which mortality is the end point. Many studies have already made successful use of the NDI including, for example, a number of cohort studies of workers at various Department of Energy facilities (21). A useful bibliography is included in the National Death Index User's Manual (20).

A recent interesting test of the NDI was provided by researchers using data from the US Nurses Health Cohort Study (22). Data were sent to the NDI on 197 known dead subjects, and 1,997subjects known to be alive. Following linkage to the NDI, and subsquent resolution of the potential links by the researchers, the sensitivity of the NDI i.e. the percentage of known dead subjects identified as such by the NDI was estimated to be 97.7percent, and the specificity i.e. the number of women known to be alive but incorrectly identified by the NDI as dead was found to be zero. These data and other similar exercises which have been reported should provide confidence in the use of the NDI as an important tool for mortality follow-up for cohort studies conducted in the U.S.

DISCUSSION

Computerized record linkage clearly has played and continues to play an important role in numerous cohort studies in Canada, the U.S. and elsewhere. The approach in conjuntion with population registries of mortality, cancer incidence and other health outcomes makes possible the use of passive follow-up which makes feasible large studies at a relatively low cost. The approach is particularly useful for retrospective cohort studies where study subjects may have been identified and lost to active follow-up many years ago.

In terms of cohort studies two problems need to be considered when record linkage is used in such studies for follow-up purposes. First, is the potential problem of out migration from coverage by the registry which is being used to identify the outcome of interest. For example, the Canadian mortality data includes death records of Canadians living in the U.S., but not Canadians who have migrated to other countries. This of course is really a problem of passive follow-up rather than of linkage per se but nevertheless is often raised in studies which have used linkage. Sometimes the problem can be addressed if migration records are available so that migrating members of the cohort can be identified. Other approaches include linking the records of the cohort to a "live registry" such as current tax files or the Social Security index in the U.S.

Although overall experience has suggested that undetected loss to follow-up is minimal in the majority of situations in which linkage has been used to date, this possibility should be borne in mind and the situation investigated before a decision is made to use passive follow up employing record linkage.

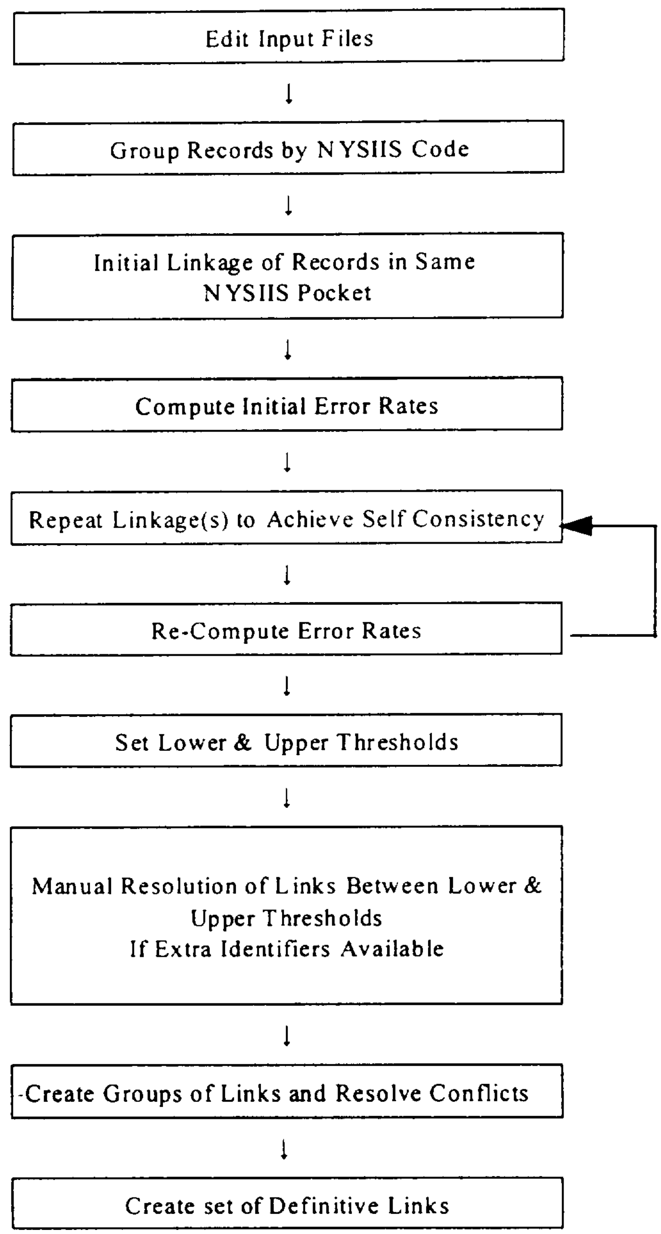
The other major potential problem arises from the probablistic nature of the linkage, so that inevitably there is some degree of uncertainty that true links may have been missed or false links introduced by the setting of a single threshold value i.e. the false negative and false positive rates may not be zero. The effect of false positives, if the rate is unrelated to exposure status, is to attenuate both risk differences and risk ratios towards the null and hence dilute any true effect. The effect of false negatives, again if the latter is unrelated to exposure, is to attenuate risk differences towards the null, but to have little impact on risk ratios, although the latter phenomenon does lead to a loss of power. Thus, a reasonable approach is to set a conservative threshold i.e. to err on the side of reducing the number of false positives if anything, and to depend for the primary analysis on relative risk rather than risk difference measures. In addition, it will be useful to repeat analyses with a variety of possible thresholds to examine how much impact this has on the corresponding relative risk estimates. In practice, in the experience of the author, this leads to relatively minor changes in relative risk estimates which is reassuring.

Overall, in terms of assessing the performance of passive follow-up in conjunction with record linkage, studies have suggested that such follow-up does as equally well as active follow-up (23). Thus, computerized record linkage should provide a reliable tool for passive follow-up in cohort studies provided careful consideration is given to the principles as described above.

One issue not addressed in this review which gives rise to concern in the application of computerized record linkage is that of confidentiality. The concern arises because linkage by definition involves personal identifiers such as names or social security numbers, and because of public perception that linkage among the data bases in which they are recorded could be done without the subject's knowledge, and with great ease. It is obviously essential therefore that strict criteria be used to maintain confidentiality. For example, the linkages to the Canadian mortality data base described earlier were all carried out in Statistics Canada under the protection of the Canadian Statistics Act, and individual records with death information were returned to researchers only with the more personal identifiers such as names removed. For prospective cohort studies there needs to be informed consent amongst study subjects to have their records linked at some future point to health outcome records. Of course this safeguard is not generally available for retrospective studies. It is reassuring to note that to date, despite the large number of studies that have utilized this technique, to the best of the author's knowledge, there has been no demonstrated breach of any individual's confidentiality.

The record linkage systems in Canada and the U.S. described above have now operated successfully for a number of years. Other countries also have linkage systems in place which permit the ready identification of health outcomes by such means including some of the Scandinavian countries (which utilize a personal identification number) and the United Kingdom (which utilizes the national health number). Further, systems are being developed in other countries including Israel and the former Soviet Union. Although some practical differences in these countries have to be addressed (e.g. the use of Cyrillic characters in the Russian language) the principles to be applied are identical to those described above with respect to the Canadian system which should continue to provide a useful model for such systems, and thus continue to facilitate the conduct of large cohort studies at reasonable cost and with high efficiency.

Figure I
Schematic of the Generalized Iterative Record Linkage System (GRLS)



References

1. Newcombe HB, Kennedy JM, Axford SJ, et al. Automatic linkage of vital records. Science 1959; 130: 315-24

2. Howe, GR, Lindsay J. A generalized iterative record linkage computer system for use in medical follow-up studies. Computers and Biomedical Research 1981; 14: 327-40

3. Newcombe HB. Handbook of record linkage: methods for health and statistical studies, administration and business. Oxford UK. Oxford University Press, 1988.

4. Miller AB, Howe GR, Sherman GJ, et al. Mortality from breast cancer after irradiation during fluoroscopic examinations in patients being treated for tuberculosis. New Engl J Med 1989; 321: 1285-89

5. Howe GR. Lung cancer mortality between 1950 and 1987 following exposure to fractionated moderate dose rate ionizing radiation in the Canadian fluoroscopy study and a comparison with lung cancer mortality in the atomic bomb survivors study. Radiat Res 1995; 142: 295-304

6. Howe GR, McLaughlin J. Breast cancer mortality between 1950 and 1987 following

exposure to fractionated moderate dose rate ionizing radiation in the Canadian fluoroscopy cohort study and a comparison with breast cancer mortality in the atomic bomb survivors study. Radiat Res 1996; 145: 694-707

7. Felligi IP, Sunter AB. A theory of record linkage. Journal Amer Stat. Assoc. 1969; 40: 1183-1210

8. Lindsay JP, Stavraky KM, Howe GR. The Canadian labor force ten percent sample study. Cancer mortality among men, 1965-1979. JOM 1993; 35: 408-14

9. Lynch BT and Arends WL. Selection of a surname coding procedure for the SRS record linkage system. Sample Survey Research Branch, Research Division, U.S. Department of Agriculture, Washington, DC 1977.

10. Howe GR. Breast cancer incidence (1975-1983) following exposure to highly fractionated ionizing radiation in the Canadian fluoroscopy cohort study. Amer J Epidemiol (in press) 1997.

11. Howe GR, Lindsay JP. A follow-up study of a ten-percent sample of the Canadian labor force 1. Cancer mortality in males, 1965-73. JNCI 1983; 70: 37-44

12. Aronson KJ, Howe GR. Utility of a surveillance system to detect associations between work and cancer among women in Canada, 1965-1991. JOM 1994; 36, 11, 1174-79

13. Howe GR, Nair RC, Newcombe HB, et al. Lung cancer mortality (1950-1980) in relation to radon daughter exposure in a cohort of workers at the Eldorado Beaverlodge uranium mine. JNCI 1986; 77: 357-62

14. Howe GR, Nair RC, Newcombe HB, et al. Lung cancer mortality (1950-1980) in a cohort of workers at the Eldorado Port Radium uranium mine: possible modification of risk by exposure rate. JNCI 1987; 79: 1255-60

15. L'Abbe KA, Howe GR, Burch JD, et al. Radon exposure , cigarette smoking, and other mining experience in the Beaverlodge uranium miners cohort. Health Physics. 1991 60: 4; 489-95

16. Howe GR, Stager RH. Risk of lung cancer mortality following exposure to radon decay products in the Beaverlodge cohort based on revised exposure estimates. Radiat Res 1996;146: 37-42

17. Lubin JG, et al. Radon and lung cancer risk: a joint analysis of 11 underground miners studies. National Institutes of Health. NIH Publ. 1994; 94-3644

18. Darby SC et al. Radon and cancers other than lung cancer in underground miners: a collaborative analysis of 11 studies. JNCI 1995; 87 (5): 378-84

19. Howe GR, Fraser D, Lindsay J, et al. Cancer mortality (1965-77) in relation to diesel fumes and coal exposure in a cohort of retired railway workers. JNCI 1983; 70: 1015-19

20. National Death Index User's Manual, National Center for Health Statistics, U.S. Department of Health & Human services, Public Health Service, Center for Disease Control, Hyattsville, Maryland, September 1990.

21. Gilbert ES, Cragle D, Wiggs LD. Updated analysis of combined mortality data for workers at the Hanford site, Oakridge National Laboratory, and Rocky Flats Weapons Plant. Radiation Research 1993; 136: 408-21

22. Rich-Edwards JW, Corsano KA, Stampfer MJ. Test of the National Death Index and Equifax Nationwide death search. Amer. J Epidemiol 1994; 140; 1016-19

23. Shannon H.S. et al. Comparison of individual follow-up and computerized record linkage using the Canadian mortality data base. Canadian J Pub. Health 1989; 80: 54-57

APPENDIX 5

REPORT ON THE NOVEMBER 14, 1997 VISIT OF DR. MASNYK AND MR. COLE TO COLUMBIA UNIVERSITY LABORATORY FACILITIES

Laboratory Tour During NCI Site Visit

The Site Review at Columbia University included a tour of the laboratory facilities of the clinical professors who are a part of the Chernobyl project. Dr. Fayter showed Dr. Masnyk and Mr. Cole the ultrasonography imaging room; Dr. Matsushima, the hematopathology laboratory where slides are processed, stained and stored; Drs. Fink and Reiss, the hematology and flow cytometry laboratories to characterize the specific type of leukemia, Dr. Greenebaum, the cytopathology four headed training microscopes and aspiration cart; and Dr. Matsushima, the surgical pathology "Gross" room where tissue specimens are received from the operating rooms, and a demonstration of the processing of a lung specimen and use of a medical voice dictation system. More detailed laboratory reports prepared by the clinical physicians and detailing the specific equipment and a description of its usage at the Columbia-Presbyterian Medical Center follow.

The NCI and Columbia team visited the cytopathology laboratory area in Babies Hospital South. The three areas include:

(1) Cytopreparatory – where specimens are accessioned into the computer, given a case number and processed as needed, centrifuged, then smeared and stained with standard Papanicolaou stain. This area handles >30,000 cervical Pap smears and 9,000 non-gynecological specimens including > 2,000 fine needle aspiration biopsy cytology cases per year. The staining solutions are stored and used under a fume hood because of carcinogenic fumes. Universal precautions – handling every case as though it is infectious with hepatitis, HIV and tuberculosis – are mandatory. It is for this reason that separate rooms are used for cytopreparation away from the cytotechnologists and clerical staff.

(2) Cytotechnologists (3 1/2 FTE) screen all the slides. The entire surface of the slide is evaluated to detect cancer cells or other neoplastic, pre-neoplastic or infected cells or infectious agents. The abnormalities are dotted with felt tip pens onto the slide's cover slipped surface. The cytotechnologist's opinion is written onto the requisition form and then entered in the computer (either by the cytotechnologist or a clerk-typist). When the case is a cervical Pap smear screening for uterine cervical cancer, only the abnormal cases and the "high risk" cases are given to the cytopathologist for review and diagnosis.

(3) In these cases the requisitions with preliminary diagnoses are delivered to the cytopathologist for final evaluation and the reports, after any necessary modifications are signed electronically after entering a unique, secret password. The report simultaneously becomes available throughout the hospital and at all off-site clinics and offices.

The group next visited the surgical pathology laboratories in the Vanderbilt Clinic. The "Gross" room where tissue specimens are received from the operating rooms, clinics and outpatient offices was visited with Dr. T. Strom, a pathology resident demonstrating the processing of a lung specimen and a medical voice dictation system for continuous speech data entry (Med Speak - IBM). The histology laboratory where tissues become sections on glass slides was visited subsequently. The automated slide labeler, 4 hour tissue processor and automated cover-slipper were seen. The special stains area is in an adjoining small enclave. It is here that the "ultra-fast Papanicolaou stain" is performed on air-dried fine needle aspiration biopsy cytology specimens (90 seconds, 10 solutions, permanent mounting). Two articles elaborating the use and benefits of this ultra-fast Papanicolaou stain, one of which is specifically about thyroid, were given to Dr. Masnyk.

The last visit was to Dr. Greenebaum's office, where a 4 "headed" Olympus BH2 microscope with 10x oculars and 2x,4x, 10x, 40x and 60x objectives is located. Use of the additional "heads" and the advantages of the ultra-fast Papanicolaou stain over the Diff/Quik or Wright Giemsa (or other Romanowsky stains) was demonstrated. The microscope's extra "heads" are used on a routine basis to teach residents and junior attending physicians advanced skills in cytopathology.

Lastly, the aspiration cart set up for attending and/or performing aspirations on hospitalized or out patients in clinics or private offices was shown to Dr. Masnyk and Mr. Cole. The cart includes a 2 "headed" binocular microscope, screw-capped Coplin jars with the 3 DiffQuick solutions I, II, III plus a 4th with water. Slides, needles and a Cameco pistol-grip syringe holder are carried along with alcohol wipes and 2 to 4 inch square gauze pads and Band-Aids. Non sterile disposable medium size exam gloves are used both to do the aspiration biopsy and the specimen preparation and staining. They are discarded and hands washed after every patient encounter.

Additionally, as a part of the facilities tour, Dr. Judith Fayter guided Dr. Masnyk and Mr. Cole through the ultrasonography area in the Department of Radiology. The imaging equipment and data storage were discussed as they relate to thyroid ultrasound. Specifically, the group was shown a standard ultrasound imaging room with an Acuson 128XP-10 including multiple transducers. For thyroid imaging, a mid-level machine would be satisfactory with 5 and 7 MHz transducers.

Image storage options include matrix and laser cameras, VCR with capability of digitizing images, or direct OD storage. Paper or thermal storage would be less reliable for long-term storage and retrieval. The discussion also included aspiration biopsy technique, specifically use of a 5/8 inch 25g needle.